# SEARCH REQUEST FORM

Requestor's
Name: _____

Serial
Number: _____

Date: _____    Phone: _____    Art Unit: _____

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevent citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevent claim(s).

## STAFF USE ONLY

Date completed: _07-29-03_

Searcher: _Beverly e 4594_

Terminal time: _22_

Elapsed time: _____

CPU time: _____

Total time: _23_

Number of Searches: _____

Number of Databases: _1_

**Search Site**

_____ STIC

_____ CM-1

_____ Pre-S

**Type of Search**

_____ N.A. Sequence

_____ A.A. Sequence

_____ Structure

_____ Bibliographic

**Vendors**

_____ IG

_____ STN

_____ Dialog

_____ APS

_____ Geninfo

_____ SDC

_____ DARC/Questel

_✓_ Other C G N

# STIC Search Report
## Biotech-Chem Library

| | |
|---|---|
| **To:** Sarvamangala Devi | **From:** Beverly Shears |
| **Location:** CM1-7E12 | **Location:** Biotech-Chem Library |
| **Art Unit:** 1645 | CM1-1E05 |
| **Tuesday, July 29, 2003** | **Phone:** 308-4994 |
| | |
| **Case Serial Number:** 10/060521 | beverly.shears@uspto.gov |

## Search Notes

OM protein - protein search, using sw model

Run on:          July 28, 2003, 15:35:01 ; Search time 98 Seconds
                                    (without alignments)
                                    806.590 Million cell updates/sec

GenCore version 5.1.6
Copyright (c) 1993 - 2003  Compugen Ltd.

Title:           US-10-060-521-2
Perfect score:   2540
Sequence:        1 MSKIELKQLSFAYDNQEVLL.........IEHDAHFMKKITDKKIVLKS 498

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5
Searched:        1107863 seqs, 158726573 residues

al number of hits satisfying chosen parameters:    2

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 75%
                 Maximum Match 100%
                 Listing first 1000 summaries

Database :       A_Geneseq_19Jun03:*

                 1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
                 2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
                 3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
                 4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
                 5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
                 6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
                 7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
                 8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
                 9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
                 10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
                 11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
                 12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
                 13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
                 14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
                 15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
                 16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
                 17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
                 18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
                 19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
                 20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
                 21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
                 22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
                 23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
                 24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | % Match | DB ID | Description |
|---|---|---|---|---|---|
| 1 | 2540 | 100.0 | 498 | 23 | ABG32864 | E. faecalis hygrom |
| 2 | 2522 | 99.3 | 498 | 22 | ABB47285 | Enterococcus faeca |

ALIGNMENTS

ID   ABG32864 standard; Protein; 498 AA.
XX
AC   ABG32864;
XX
DT   20-NOV-2002  (first entry)
XX
DE   E. faecalis hygromycin A resistance protein, harA.
XX
KW   HarA; enzyme; hygromycin A resistance; bacterial infection; sepsis;
KW   septic shock; bacterial endocarditis; prostatitis; NTPase;
KW   staphylococcal infection; pneumococcal infection.
XX
OS   Enterococcus faecalis strain OG1X.
XX
PN   EP1227324-A2.
XX
PD   31-JUL-2002.
XX
PF   29-JAN-2002; 2002EP-0250612.
XX
PR   30-JAN-2001; 2001US-265034P.
XX
PA   (PFIZ ) PFIZER PROD INC.
XX
PI   Baima ET, Mueller JP;
XX
DR   WPI; 2002-676579/73.
DR   N-PSDB; ABS53679.
XX
PT   Novel hygromycin A resistance polypeptide derived from Enterococcus
PT   faecalis and Bacillus subtilis useful for screening compounds which
PT   binds to the polypeptide, and as medicament for treating drug resistant
PT   infection -
XX
PS   Claim 12; Fig 2; 43pp; English.
XX
CC   The invention relates to a recombinant hygromycin A resistance (harA)
CC   polypeptide comprising a sequence which is at least 75% identical to
CC   the harA protein form Enterococcus faecalis and Bacillus subtilis.
CC   Also included are: (1) use of a hygromycin A-resistant strain of
CC   Enterococcus faecalis or Bacillus subtilis to determine whether an
CC   antibacterial agent is effective in treating organisms which exhibit
CC   harA-mediated drug resistance; (2) the nucleic acid encoding the harA
CC   polypeptides and (3) harA expression vectors and host cells.
CC   HarA protein or nucleic acid is useful in the screening/manufacture of
CC   a compound for use as a medicament for the treatment of a drug resistant
CC   infection. HarA nucleic acid is useful to identify an organism containing
CC   a harA gene. HarA protein is useful to screen for E. faecalis antibodies
CC   in blood or serum and samples from animals, or as antigens to raise
CC   antibodies, which antibodies are useful as diagnostic reagents. The
CC   identified compounds are useful in the treatment of enterococcal
CC   infections (e.g. sepsis, septic shock, bacterial endocarditis and
CC   prostatitis) as well as staphylococcal and pneumococcal infections.
CC   The nucleic acid is useful as a hybridisation probe for
CC   RNA, cDNA and genomic DNA to isolate full-length cDNAs and genomic
CC   clones encoding harA polypeptides and to isolate cDNA and genomic
CC   clones of other genes that have a high sequence similarity to harA
CC   gene. HarA has NTPase activity. The present sequence is the B. faecalis
CC   harA protein.
XX
SQ   Sequence  498 AA;

Query Match       100.0%;  Score 2540;  DB 23;  Length 498;
Best Local Similarity   100.0%;
Matches  498;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy          1 MSKIELKQLSFAYDNQEVLLFDQANITMDTNWKLGLIGRNGRGKTTLLRLLQKOLDYQGB 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MSKIELKQLSFAYDNQEVLLFDQANITMDTNWKLGLIGRNGRGKTTLLRLLQKOLDYQGB 60

Qy         61 ILHQVDFVFPPQTVABEQQLTYYVLQBVTSFBQWELRRELTLLNVDPEVLWRPFSSLSGG 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 ILHQVDFVFPPQTVABEQQLTYYVLQBVTSFBQWELRRELTLLNVDPEVLWRPFSSLSGG 120

CC further identifying MDR efflux pumps that may be used as drug targets to
CC increase the sensitivity of cells to antibacterial agents. Cells
CC comprising the identified pumps may be used to screen for potential
CC blockers or inhibitors of MDR pump function or gene expression.
XX
SQ Sequence 498 AA;

```
Query Match          99.3%;  Score 2522;  DB 22;  Length 498;
Best Local Similarity   99.2%;
Matches  494;  Conservative   2;  Mismatches   2;  Indels   0;  Gaps   0;
                                             Pred. No. 5.5e-189;

Qy      1 MSKIELKQLSPAYDNQBVLLFDQANITMDTNWKLGLIIGRNGRGKTTLLLRLLQKQLDYQGB 60
          |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
Db      1 MSKIELKQLSPAYDNQBALLPDQANITMDTNWKLGLIIGRNGRGKTTLLLRLLQKQLDYQGB 60

Qy     61 ILHQVDFVYFPQTVABEQQLTYYVLQBVTSFEQWELERERLTLLNVDPEVLWRPFSSLSGG 120
          |||||||||:||||||||::|||||||||||||||||||||||||||||||||||||||||
Db     61 ILHQVDFVYFPQTVABEQQLTYYVLQBVTSFBQWCLERBLTLLNVDPEVLWRPFSSLSGG 120

Qy    121 EKTKVLLGLLFIBENAFPLIDBPTNHLDLAGRQQVABYLKKKGHGPILVSHDRAPVDEVV 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 EKTKVLLGLLFIBENAFPLIDBPTNHLDLAGRQQVABYLKKKGHGPILVSHDRAPVDEVV 180

Qy    181 DHILAIEKSQLTLYQGNFSIYBEQKKLRDAFBLAENEKIKKEVNRLKETARKKAEWSMNR 240
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 DHILAIEKSQLTLYQGNFSIYBEQKKLRDAFBLAENEKIKKEVNRLKETARKKAEWSMNR 240

Qy    241 EGDKYGNAKEKGSGAIFDTGAIGARAARVMKRSKHIQORABTQLABKRKLLKDLEYIDPL 300
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 EGDKYGNAKEKGSGAIFDTGAIGARAARVMKRSKHIQORABTQLABKRKLLKDLEYIDSL 300

Qy    301 SMDYQPTHHKTLLTVEBLRLGYBKNWLFAPLSFSINAGBIVGITGKNGSGKSSLIQYLLD 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 SMDYQPTHHKTLLTVEBLRLGYBKNWLFAPLSFSINAGBIVGITGKNGSGKSSLIQYLLD 360

Qy    361 NFSGDSEGBATLAHQLTISYVRQDYEDNQGTLSEFABKRNQLDYTQFLNNLRKLGMBRAVP 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 NFSGDSEGBATLAHQLTISYVRQDYEDNQGTLSEFABKRNQLDYTQFLNNLRKLGMBRAVP 420

Qy    421 TNRIBQMSMGQRKKVBVAKSLSQSAELYIWDBPLNYLDVFNHQQLBALILSVKPAMLVIB 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TNRIBQMSMGQRKKVBVAKSLSQSAELYIWDBPLNYLDVFNHQQLBALILSVKPAMLVIB 480

Qy    481 HDAHFMKKCITDKKIVLKS 498
          |||||||||||||||||||
Db    481 HDAHFMKKCITDKKIVLKS 498
```

RESULT 2
ABB47285
ID ABB47285 standard; Protein; 498 AA.
XX
AC ABB47285;
XX
DT 31-JAN-2002 (first entry)
XX
DE Enterococcus faecalis polypeptide Abc23.
XX
KW MDR; efflux pump; multidrug resistance; antibacterial; drug target.
XX
OS Enterococcus faecalis.
XX

PT The invention relates to determining whether a candidate nucleotide
PT (ABA82938-ABA82971) or polypeptide (ABB47263-ABB47286) encodes/functions
PT as a multidrug resistance (MDR) efflux pump comprising, searching a
PT database for sequences high identity known MDR efflux pumps and then
PT deleting/mutating an identified region of the DNA in a bacterial cell and
PT determining whether the bacterial cell exhibits increased or decreased
PT sensitivity to an antibacterial agent The identified pumps are useful for

Claim 10; Fig 26; 139pp; English.

Determining whether a candidate nucleotide or polypeptide
encodes/functions as a multidrug resistance (MDR) efflux pump comprises
searching a database of nucleotide/polypeptide sequences for those with
high identity to known MDR pumps -

```
Qy    301 SMDYQPTHHKTLLTVEBLRLGYBKNWLFAPLSFSINAGBIVGITGKNGSGKSSLIQYLLD 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 SMDYQPTHHKTLLTVEBLRLGYBKNWLFAPLSFSINAGBIVGITGKNGSGKSSLIQYLLD 360

Qy    361 NFSGDSEGBATLAHQLTISYVRQDYEDNQGTLSEFABKRNQLDYTQFLNNLRKLGMBRAVP 420
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 NFSGDSEGBATLAHQLTISYVRQDYEDNQGTLSEFABKRNQLDYTQFLNNLRKLGMBRAVP 420

Qy    421 TNRIBQMSMGQRKKVBVAKSLSQSAELYIWDBPLNYLDVFNHQQLBALILSVKPAMLVIB 480
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TNRIBQMSMGQRKKVBVAKSLSQSAELYIWDBPLNYLDVFNHQQLBALILSVKPAMLVIB 480

Qy    481 HDAHFMKKCITDKKIVLKS 498
          |||||||||||||||||||
Db    481 HDAHFMKKCITDKKIVLKS 498
```

Search completed: July 28, 2003, 15:39:59
Job time : 99 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          July 28, 2003, 15:40:06 ; Search time 30 Seconds
                                            (without alignments)
                                            702.360 Million cell updates/sec

Title:
Perfect score:   2540
Sequence:        1 MSKIELKQLSFAYDNQBVLL..........IEHDAHFMKITDKIVLKS 498

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        328717 seqs, 42310858 residues

al number of hits satisfying chosen parameters:   0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 75%
                 Maximum Match 100%
                 Listing first 1000 summaries

Database :       Issued_Patents_AA:*
                 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
                 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
                 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
                 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
                 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                         SUMMARIES

Result         Query
No.    Score   Match Length DB   ID                    Description
-----  -----   ------------ --   --                    -----------

o matches found

Search completed: July 28, 2003, 15:50:53
Job time : 30 secs

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: July 28, 2003, 15:38:06 ; Search time 53 Seconds
(without alignments)
1115.894 Million cell updates/sec

Title: US-10-060-521-2
Perfect score: 2540
Sequence: 1 MSKIELKQLSFAYDNQBVLL..............IBHDAHFMKITDKKIVLKS 498

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 451899 seqs, 118759770 residues

al number of hits satisfying chosen parameters: 0

Post-processing: Minimum Match 75%
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Maximum Match 100%
Listing first 1000 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptcodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptcodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptcodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptcodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptcodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptcodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptcodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptcodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptcodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptcodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptcodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptcodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptcodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptcodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptcodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptcodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptcodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptcodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Match Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| ------ | ----- | ------------ | ------ | -- | -- | ----------- |

No matches found

Search completed: July 28, 2003, 15:48:26
Job time : 53 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          July 28, 2003, 15:36:25 ; Search time 52 Seconds
                                           (without alignments)
                                           921.000 Million cell updates/sec

Title:
Perfect score:   2540
Sequence:        1 MSKELKQLSFAYDNQBVLL.............IEHDAHFMKKITDKKIVLKS 498

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        283308 seqs, 96168682 residues

al number of hits satisfying chosen parameters:    0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 75%
                 Maximum Match 100%
                 Listing first 1000 summaries

Database :       PIR_76:*
                 1:  pir1:*
                 2:  pir2:*
                 3:  pir3:*
                 4:  pir4:*

                 Pred. No. is the number of results predicted by chance to have a
                 score greater than or equal to the score of the result being printed,
                 and is derived by analysis of the total score distribution.

                                SUMMARIES

Result           Query
  No.    Score   Match Length DB    ID              Description
-------  -----   ------------ --    --              -----------

No matches found

rch completed: July 28, 2003, 15:40:58
time : 52 secs

```
                            GenCore version 5.1.6
                      Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - protein search, using sw model

Run on:        July 28, 2003, 15:35:01 ; Search time 25 Seconds
                                          (without alignments)
                                          936.772 Million cell updates/sec

Title:
Perfect score:   2540
Sequence:      1 MSKIELKQLSFAYDNQEVLL............IEHDAHFMKKITDKKIVLKS 498

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        127863 seqs, 47026705 residues

al number of hits satisfying chosen parameters:    0

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 75%
                 Maximum Match 100%
                 Listing first 1000 summaries

Database :    SwissProt_41:*

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                            SUMMARIES

              %
Result        Query
  No.   Score Match Length DB   ID            Description
------- ----- ----- ------ ---- ----------    -----------

No matches found

Search completed: July 28, 2003, 15:38:21
Job time : 33 secs
```

OM protein - protein search, using sw model

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

Run on: July 28, 2003, 15:38:26 ; Search time 103 Seconds
(without alignments)
1247.672 Million cell updates/sec

Title: US-10-060-521-2
Perfect score: 2540
Sequence: 1 MSKIELKQLSFAYDNQEVLL..........IEHDAHFMKKITDKIVLKS 498

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

al number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 75%
Maximum Match 100%
Listing first 1000 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | DB | ID | Description |
|---|---|---|---|---|---|
| ------- | ------- | ------------ | --- | --- | ------------ |

No matches found

Search completed: July 28, 2003, 15:50:22
Job time : 110 secs